Trend test and change-point detection for the annual discharge series of the Yangtze River at the Yichang hydrological station

LIHUA XIONG & SHENGLIAN GUO
State Key Laboratory of Water Resources and Hydropower Engineering Sciences, Wuhan University, Wuhan 430072, China
lhxiong@public.wh.hb.cn

Abstract The trend test and change-point analysis have been carried out on several annual discharge series of the Yangtze River at the Yichang hydrological station, including the discharge series of the annual maximum, annual minimum and annual mean during the period 1882–2001. The results of both the Mann-Kendall and Spearman's rho trend tests show that, at the 5% significance level, the annual maximum flood series did not have any statistically significant trend, but the annual minimum flow series and the annual mean discharge series exhibited a sign of decreasing trend. For single change-point detection, a Bayesian model was established to study the abrupt change in the mean levels of the time series. Given the observed hydrological data, the Bayesian model can estimate the posterior probability distribution of each change-point location by using the Monte Carlo Markov Chain (MCMC) sampling method. The results of the Bayesian model show that, during the past 120 years, the mean levels of both the annual minimum discharge series and the annual mean discharge series have decreased by 8% and 6% respectively. Further analysis suggested that, for the annual minimum discharge series and the annual mean discharge series, both the trend term and the abrupt change term are very closely related and very hard to distinguish from each other.

Key words trend test; change-point analysis; Bayesian inference; posterior probability distribution; annual discharge series; Yangtze River

Test de tendance et détection de rupture appliqués aux séries de débit annuel du fleuve Yangtze à la station hydrologique de Yichang

Résumé Des tests de tendance et des analyses de rupture ont été appliqués à plusieurs séries annuelles (séries des débits annuels maximums, minimums et moyens) du fleuve Yangtze, au niveau de la station de jaugeage de Yichang, sur la période 1882–2001. Les résultats des tests de tendance de Mann-Kendall d'une part et rho de Spearman d'autre part montrent que, pour un niveau de signification inférieur à 5%, la série des débits maximums annuels ne présente pas de tendance significative, mais que celles des débits minimums et moyens annuels présentent une tendance décroissante. Pour la détection d'un point de rupture simple, un modèle bayésien a été établi afin d'étudier le changement brutal du niveau moyen des séries. Vu les données hydrologiques observées, le modèle bayésien permet d'estimer la distribution de probabilité *a posteriori* de la localisation du point de rupture par une méthode d'échantillonnage aléatoire. Les résultats du modèle bayésien montrent que, lors des 120 dernières années, les niveaux moyens des séries des débits minimum et moyens annuels ont respectivement diminué de 8% et 6%. Une analyse plus approfondie suggère que, pour les séries des débits minimums et moyens annuels, les termes de tendance et de rupture sont étroitement liés et qu'il est très difficile de les distinguer.

Mots clefs test de tendance; analyse de rupture; inférence bayésienne; distribution de probabilité *a posteriori*; séries de débit annuel; fleuve Yangtze

INTRODUCTION

Since its actual commencement in 1993 (formally on 14 December 1994), the construction of the Three Gorges Project, which is located about 40 km upstream of

Open for discussion until 1 August 2004

Yichang City in Hubei Province, central China, has progressed very smoothly. According to schedule, the first generator set was installed and began to operate by the end of 2003; the whole project will be completed in 2009. Three major functions of the Three Gorges Project are flood control, navigation and hydropower generation. With its total installed capacity of 18 200 MW, the Three Gorges Hydropower Station can generate 84.7 billion kWh a year and is the largest hydropower plant in the world. As the planning, design, construction and operation of the Three Gorges Project are based mainly on hydrological data, typically the discharge, at the Yichang hydrological station, it is essential and important to investigate the properties of the long time series of these data.

In water resources planning studies, the hydrological time series used are supposed to meet a set of ideal conditions, such as being consistent and trend-free (Adeloye & Montaseri, 2002). Consistency implies that all the collected data belong to the same statistical population, while trend-free means that there is no significant correlation between the observations and time. However, these assumptions may not hold in reality. As reported in many published studies, the hydrological data series from many regions demonstrate significant non-consistency or non-stationarity, due to the influence of climate change and/or large-scale human activities on the water resources systems. For example, a number of recent studies in the USA have detected the presence of trends in streamflow data (Smith & Richman, 1993; Pupacko, 1993; Changnon & Kunkel, 1995; Lettenmaier *et al.*, 1994; Lins & Slack, 1999; Olsen *et al.*, 1999; Douglas *et al.*, 2000). Some other literature reported the non-consistency of hydrological time series and found that there were one or more change points present within the observed data series (e.g. Paturel *et al.*, 1997; Perreault *et al.*, 1999, 2000a,b; Servat *et al.*, 1997).

This paper investigates whether the discharge records of the Yangtze River at the Yichang hydrological station exhibit evidence of change—either gradual change (trend) or abrupt change (jump)—during the period 1882–2001. The methods employed for trend test are the two widely used non-parametric methods, i.e. the Mann-Kendall test (Mann, 1945; Kendall, 1975; Meddis, 1984; Hirsch & Slack, 1984; Burn, 1994; Burn & Hag Elnur, 2002; Gan, 1998; Zhang *et al.*, 2001; Yue & Wang, 2002; Yue *et al.*, 2002a) and Spearman's rho test (Lehmann, 1975; Sneyers, 1990; Meddis, 1984; Crawshaw & Chambers, 1990; Yue *et al.*, 2002a). To identify whether there is a jump term in the components of the time series, the Bayesian approach, rather than other methods such as the Hubert segmentation method (Hubert *et al.*, 1989) and the U Buishand statistic (Buishand, 1984), was used. This has an advantage of estimating the posterior probability distribution of the change-point location (Chib, 1998; Perreault *et al.*, 2000a,b; Lavielle & Lebarbier, 2001).

PRELIMINARY DATA ANALYSIS

The Yichang hydrological station is located about 40 km downstream of the Three Gorges Project and has discharge records dating back to 1882. The series of annual maximum flood Q_{max}, the annual minimum flow Q_{min} and the annual mean discharge Q_{ave} during the period 1882–2001 (see Figs 1, 2 and 3 respectively) were extracted from the daily discharge records and used to represent the long-term hydrological characteristics at the Three Gorges Project site. As the methods employed in this paper

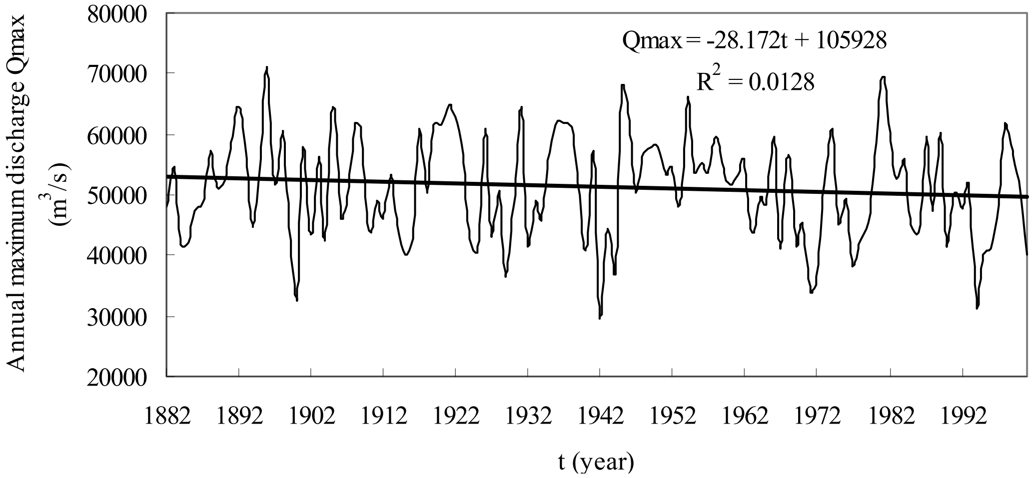

Fig. 1 Annual maximum discharge series of the Yangtze River at the Yichang hydrological station.

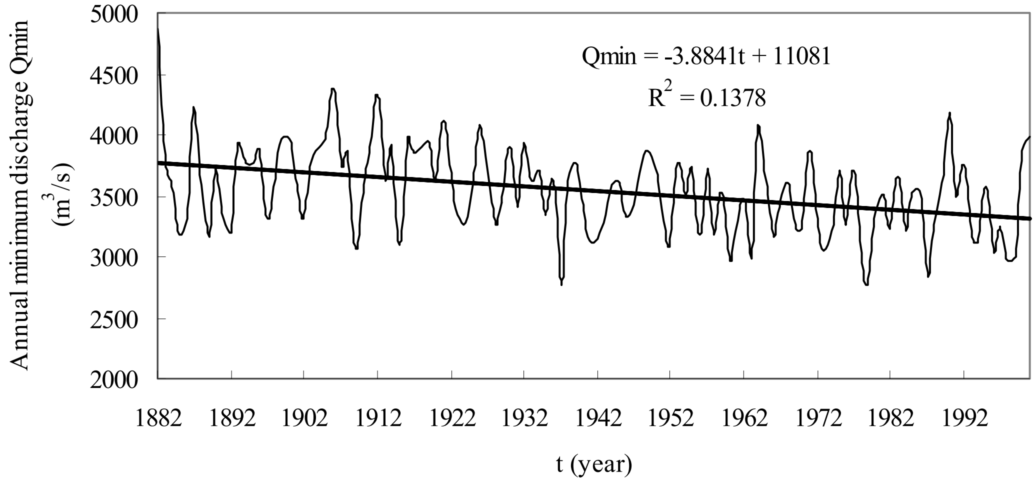

Fig. 2 Annual minimum discharge series of the Yangtze River at the Yichang hydrological station.

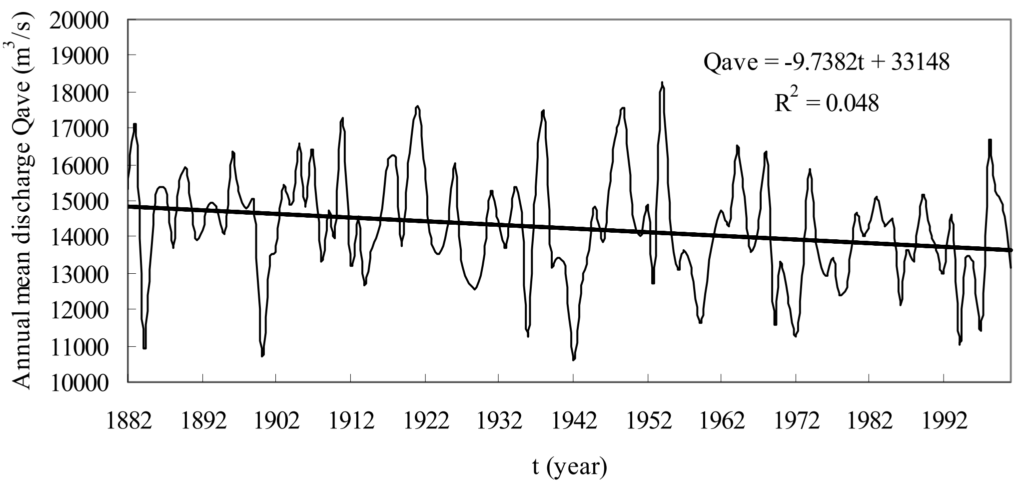

Fig. 3 Annual mean discharge series of the Yangtze River at the Yichang hydrological station.

for trend test and change-point analysis are based on the assumption of independence and normal distribution of the data, it was checked first whether the three annual series met these assumptions.

The independence test was carried out mainly by examining the autocorrelation coefficients of the time series. When the absolute values of the autocorrelation coefficients of different lag times calculated for a time series consisting of n observations are not larger than the typical critical value, i.e. $1.96/\sqrt{n}$ corresponding to the 5% significance level (Douglas *et al.*, 2000), the observations in this time series can be accepted as being independent from each other. According to the calculated autocorrelation coefficients of lag-1 to lag-10 for each annual series, the observations in that series can be accepted as being independent at the 5% significance level. For example, the lag-1 autocorrelation coefficients for the annual maximum, annual minimum and annual mean series are 0.151, 0.170 and 0.179, respectively, and all are smaller than $1.96/\sqrt{n} = 1.96/\sqrt{120} = 0.181$.

The normal probability plot was used to test for the normality of the three annual series. Of the three time series, the annual mean series fits the normal distribution best (Fig. 4), the annual maximum series second, and the annual minimum series third. This conclusion was supported by the value of the coefficient of skewness of the observations, which is 0.06 for the annual mean series, –0.11 for the annual maximum series, and 0.30 for the annual minimum series. According to the above analysis, the observations in both the annual mean series and the annual maximum series were assumed to follow the normal distribution. Although the annual minimum series has a certain

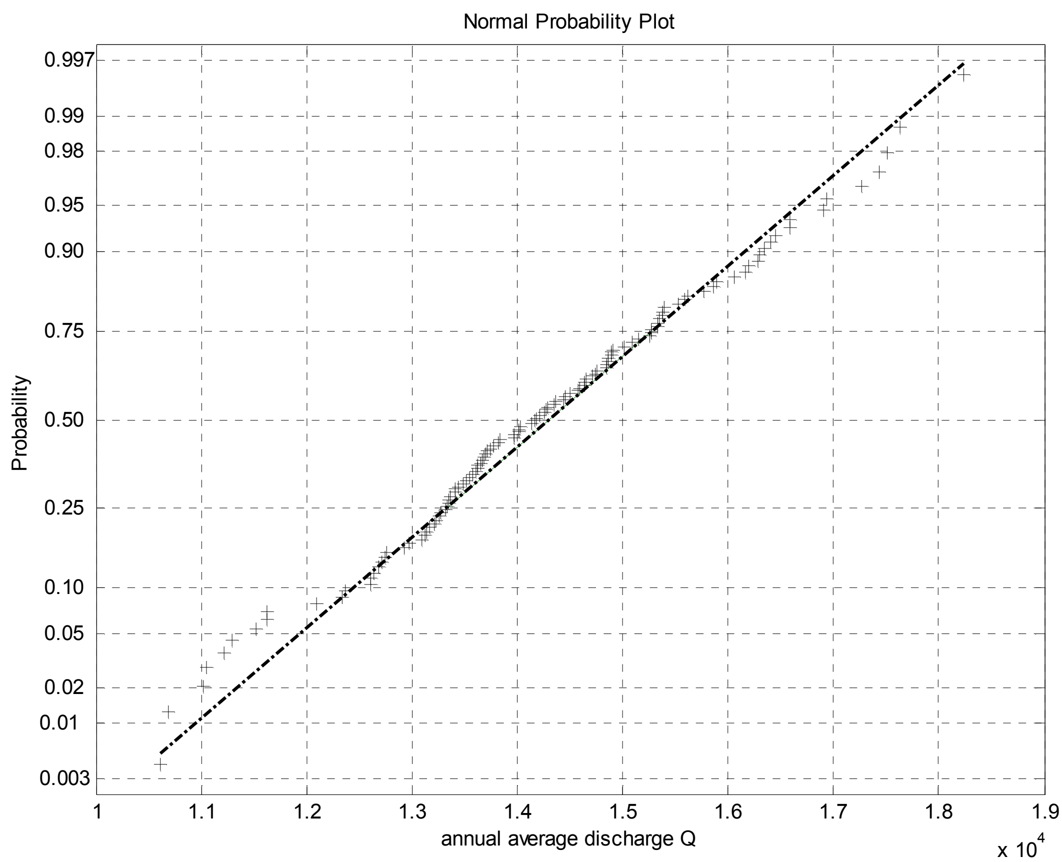

Fig. 4 The normal probability plot for the annual mean discharge series.

degree of skewness, it was also assumed to follow the normal distribution, like the other two annual series. This assumption should be reasonable considering that the discharge at the Yichang hydrological station is the summation of the numerous inflows from the vast upstream region. It must be noted that the Pearson-III probability distribution is often adopted to fit the annual discharge series in China, as it is more flexible than the normal distribution (Zhan & Ye, 2000).

TREND TEST

Two non-parametric methods, i.e. the Mann-Kendall (MK) test and Spearman's rho (SR) test were used to detect trends in the three annual series. In addition, linear regression was used to quantitatively describe the possible linear trend in the time series. For the linear regression method being used to detect the increasing or decreasing linear trend in the hydrological time series, the test statistic T is defined as:

$$T = \hat{b} / se(\hat{b}) \tag{1}$$

where $\hat{b}$ is the estimated slope of the regression line between the observation values and time and $se(\hat{b})$ represents the standard error of $\hat{b}$. For the null hypothesis that the observation values and time are not linearly related, the test statistic T has a t distribution with $n - 2$ degrees of freedom, where n is the length of the sample (Utts & Heckard, 2002). Given the significance level α for the two-tailed test, the null hypothesis is accepted if $|T| < t_{\alpha/2}(n - 2)$; otherwise, the null hypothesis is rejected.

For the t-test method to be used in the trend analysis, it requires that the observations in a time series be independent, as the existence of autocorrelation in time series will, to some degree, affect the test results (Yue *et al.*, 2002b). It also requires that the observations are normally distributed. According to the preliminary data analysis, these three annual series generally meet the conditions of independence and normality. The results of the above three trend tests are listed in Table 1. From Table 1, it is found that the results of the two non-parametric test statistics Z_{MK} and Z_{SR} are in agreement. At a 5% significance level, the null hypothesis of no trend should be accepted for the 120-year annual maximum flood series $Q_{\max}$, but it should be rejected for both the 120-year annual minimum flow series $Q_{\min}$ and the 120-year annual mean discharge series Q_{ave}. Comparing the $Q_{\min}$ and Q_{ave} series, the former exhibited much stronger evidence of trend, as its p values corresponding to both Z_{MK} and Z_{SR} are 0.000.

The estimated linear regression lines for the three hydrological series are also plotted in Figs 1, 2 and 3 respectively; the values of test statistic T for the regression slope $\hat{b}$ are listed in Table 1. In Figs 1–3, R^2 refers to the determination coefficient of the linear regression, while R is the correlation coefficient. At the 5% significance level, the linear relationship is not significant for the $Q_{\max}$ series, as $|T| = 1.235 < t_{\alpha/2}(n - 2) = t_{0.05/2}(120 - 2) = 2.014$; for the $Q_{\min}$ and Q_{ave} series, there are significant (negative) linear relationships, as their values of $|T|$, i.e. 4.342 and 2.439 respectively, are larger than 2.014. However, the $Q_{\min}$ series, shows the strongest sign of auto-correlation.

From the above trend analysis, it is noticed that the series of the annual minimum flow $Q_{\min}$ and the annual mean discharge Q_{ave} of the Yangtze River for the period

Table 1 Results of the trend test for the series of annual maximum, annual minimum and annual mean.

Time series	Mann-Kendall test		Spearman's rho test		Linear regression	
	Z_{MK}	p value	Z_{SR}	p value	$\hat{b}$	T
$Q_{\max}$	1.268	0.205	1.216	0.224	–28.172	–1.235
$Q_{\min}$	4.075	0.000	3.973	0.000	–3.884	–4.342
Q_{ave}	2.720	0.006	2.739	0.006	–9.738	–2.439

1882–2001 at the Yichang hydrological station exhibit signs that the long-term discharge of the Yangtze River from the upper stream is more or less decreasing with time. Qin *et al.* (1993) pointed out that this decreasing trend is mainly caused by the increasing use of water for agricultural and industrial development in the upstream regions of the Yangtze River, as there have not been any significant changes in the precipitation pattern or amount. For the $Q_{\max}$ series, this influence of consumptive water use can be negligible, since the quantity of water used can be regarded as negligible during the maximum flood period. However, as the magnitude of the discharge decreases, this impact of influence is increasing as the quantity of water used becomes proportionally more significant during the corresponding flow period, especially during dry periods. This can explain, to some extent, why in the above analysis the decreasing trend in the Q_{ave} series is stronger than in the $Q_{\max}$ series but much weaker than in the $Q_{\min}$ series.

CHANGE-POINT ANALYSIS

Hydrological processes are always under the influence of climate and human activity. Some particular climatic phenomena such as El Niño, as well as all kinds of large-scale water resources development projects, may alter hydrological processes suddenly and lead to abrupt change in the hydrological time series. Thus, studying whether there are any abrupt changes in the time series or not, and identifying the locations of change points, if there are any, are both essential for checking the stationarity and consistency assumptions. As discussed above, many methods are applied for change-point analysis, from the non-parametric methods (Servat *et al.*, 1997), the fuzzy method (Yu *et al.*, 2001) to all kinds of Bayesian approaches (Chernoff & Zacks, 1963; Lee & Heghinian, 1977; Berger, 1985; Paturel *et al.*, 1997; Perreault *et al.*, 1999, 2000a,b). In this study, the Bayesian approach was adopted to detect change in the mean levels of both the annual minimum flow series and the annual mean discharge series of the Yangtze River at the Yichang hydrological station; the Bayesian method has the advantage of making inferences on the posterior distribution with respect to change-point location.

The Bayesian model for a single change in the mean level

In the literature, the Bayesian model for a single change in the mean levels of the time series is the most widely used (Chernoff & Zacks, 1963; Lee & Heghinian, 1977; Berger, 1985; Kotz & Wu, 2000; Perreault *et al.*, 2000a). Consider a hydrological time series $X = \{x_1, x_2, ..., x_n\}$, whose first and second segments fluctuate, due to some exogenous factors, around different mean levels, respectively, μ_a and μ_b, but with the same variance σ^2. Denote the change-point location by $k(1 \le k < n)$, and the two segments divided by the change point as follows:

$$X_k = \{x_1, x_2, \cdots, x_k\} \qquad X_{k+1} = \{x_{k+1}, x_{k+2}, \cdots, x_n\} \tag{2}$$

In the following derivation, it is required that the time series follow the normal distribution. In the event that the hydrological data fail the normality test, the Box-Cox transformation (Box & Jenkins, 1976) must be implemented to transform the original data series into a new series with a normal distribution. This was unnecessary here, as it was concluded that both the annual minimum flow series and the annual mean discharge series of the Yangtze River could be assumed to be normally distributed.

The normal distribution functions for the random variable in each segment are denoted by:

$$x_i \sim N(\mu_a, \sigma^2) \qquad i = 1, 2, ..., k \tag{3}$$

$$x_i \sim N(\mu_b, \sigma^2) \qquad i = k+1, ..., n \tag{4}$$

When one is interested only in the shift in the mean level of the time series, one can treat the mean level of the time series as the random variable following a certain distribution. Typically, the prior distributions of both μ_a and μ_b are also assumed to be the same normal distribution, denoted by:

$$\mu_a \sim N(\mu_0, \sigma_0^2) \qquad \mu_b \sim N(\mu_0, \sigma_0^2) \tag{5}$$

When the variance σ_0^2 is large enough, the normal distributions in equation (5) will approach the non-informative prior distribution.

The variance σ^2 of the hydrological series can be regarded as a constant (Lee & Heghinian, 1977) and estimated by the sample of the hydrological series, i.e.:

$$\sigma^2 \approx \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{(n-1)} \qquad \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{6}$$

It is noted in some Bayesian models based on normal distributions, that the variance σ^2 is assumed to be an unknown random variable and follows some specific kind of distribution, such as the widely used inverted Gamma distribution, for the convenience of the mathematics (Berger, 1985; Perrault *et al.*, 2000a).

After a sample of time series X is observed, then the posterior distribution of the mean levels μ_a and μ_b can be determined according to the Bayes theorem:

$$\mu_a | X_k \sim N(\mu_a^*, \sigma_a^{*2}) \tag{7}$$

$$\mu_a^* = \frac{n^* \cdot \mu_0 + \sum_{i=1}^{k} x_i}{n^* + k} \qquad \sigma_a^{*2} = \frac{\sigma^2}{n^* + k} \qquad n^* = \frac{\sigma^2}{\sigma_0^2} \tag{8}$$

$$\mu_b | X^{k+1} \sim N(\mu_b^*, \sigma_b^{*2}) \tag{9}$$

$$\mu_b^* = \frac{n^* \cdot \mu_0 + \sum_{i=k+1}^{n} x_i}{n^* + (n-k)}, \quad \sigma_b^{*2} = \frac{\sigma^2}{n^* + (n-k)}, \quad n^* = \frac{\sigma^2}{\sigma_0^2} \tag{10}$$

In the following steps, the posterior distribution of the change-point location k is derived. The likelihood function resulting from n observations of $X = \{x_1, x_2, ..., x_n\}$ generated by equations (3) and (4) can be written as:

$$p(X|k,\mu_a,\mu_b) = \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x_i - \mu_a)^2}{2\sigma^2}\right] \prod_{i=k+1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x_i - \mu_b)^2}{2\sigma^2}\right] \tag{11}$$

Using the Bayes theorem, the posterior distribution of the change-point location k is derived as:

$$p(k|X,\mu_a,\mu_b) = \frac{p(X|k,\mu_a,\mu_b) \cdot p(k)}{\sum_{j=1}^{n-1} p(X|j,\mu_a,\mu_b) \cdot p(j)} \tag{12}$$

where $p(j)$ represents the prior distribution of the change-point location k, which is often assumed to be a uniform distribution, i.e. $p(j) = 1/(n-1)$, $j = 1, ..., n-1$. Under such an assumption, equation (12) reduces to:

$$p(k|X,\mu_a,\mu_b) = \frac{p(X|k,\mu_a,\mu_b)}{\sum_{j=1}^{n-1} p(X|j,\mu_a,\mu_b)} \tag{13}$$

The full conditional distribution of k, i.e. $p(k|X)$, may not be expressed in a simple form, but it can be estimated by some Markov Chain Monte Carlo (MCMC) approaches (Smith & Roberts, 1993; Chib, 1996, 1998), especially the Metropolis-Hasting algorithm (Metropolis *et al.*, 1953; Kuczera & Parent, 1998; Xiong & O'Connor, 2000).

Results of change-point analysis

For the annual minimum flow series Q_{min} of the Yangtze River at the Yichang hydrological station, the assumed prior and the derived posterior distributions of the change-point location are plotted in Fig. 5. It is found from Fig. 5 that the maximum posterior probability (0.11) was in 1934, which means that the mean level of the Q_{min} series changed around this year. The mean level of the Q_{min} series is 3707 m^3 s^{-1} during the period 1882–1934 and 3408 m^3 s^{-1} during the period 1935–2001 (Fig. 6). If 1934 is taken as a change point, the mean level of the Q_{min} series decreases by nearly 300 m^3 s^{-1} or 8%. Using a two-sample (two-tailed) t-test method to test the difference between the two independent samples' mean levels (Crawshaw & Chambers, 1990; Utts & Heckard, 2002), the null hypothesis that the two samples separated by the 1934 change point have the same mean level was rejected at the 5% significance level. That is, the decrease of the mean level of the Q_{min} series is statistically significant.

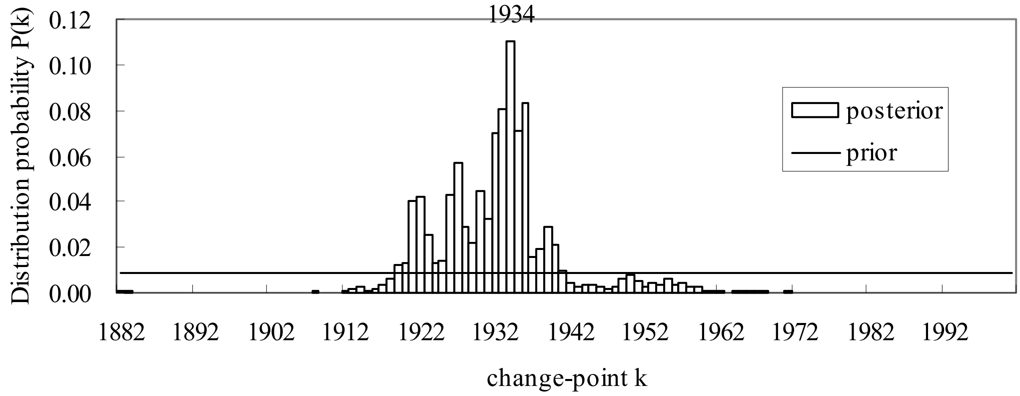

Fig. 5 The prior and posterior probability distribution of change-point location over the annual minimum flow series (1882–2001) of the Yangtze River.

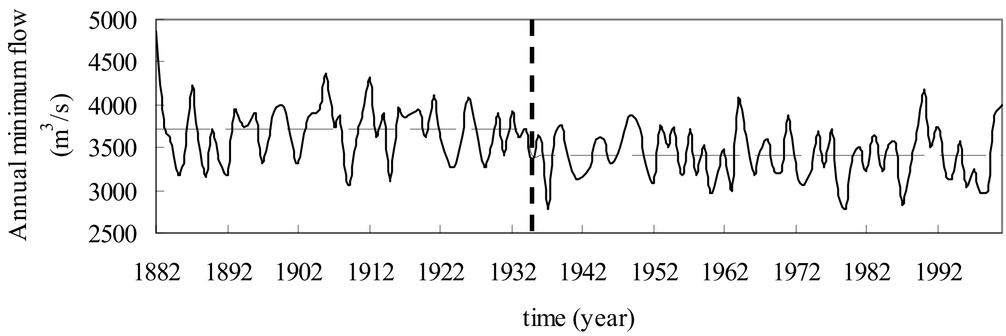

Fig. 6 The change of the mean levels over the annual minimum flow series (1882–2001) of the Yangtze River.

In order to further test whether or not the year 1993 is a potential change point for the annual minimum series, after the detected change point of 1934, the annual minimum series of 1935–2001 was divided into two segments: one of 1935–1993 with a mean value of 3416 $m^3 s^{-1}$, and the second of 1994–2001 with a mean value of 3349 $m^3 s^{-1}$. Using the two-sample (two-tailed) *t*-test method, the null hypothesis that the two segments separated by 1993 have the same mean level was not rejected at the 5% significance level. That is, there is no evidence that the year 1993 is a change point in the annual minimum series of 1935–2001.

For the annual mean discharge series Q_{ave}, the prior and posterior distributions of the change-point location are plotted in Fig. 7. Figure 7 shows that the maximum posterior probability (0.052) occurred in 1968 (change point for the mean level of the Q_{ave} series). The mean level of the Q_{ave} series is 14 486 $m^3 s^{-1}$ for the period 1882–1968 and 13 597 $m^3 s^{-1}$ for 1969–2001 (Fig. 8); i.e. the mean level of the Q_{ave} series decreased by nearly 900 $m^3 s^{-1}$ or 6%. Using the two-sample (two-tailed) *t*-test method to test the difference between the two samples' mean levels, the null hypothesis that the two samples separated by the detected change point 1968 have the same mean level was rejected at the 5% significance level. That is, the decrease of the mean level of the Q_{ave} series is statistically significant. The change-point location found for the Q_{ave} series of 1882–2001 is in agreement with the conclusion of Qin *et al.* (1993), who identified the change-point location as 1968 for the Q_{ave} series of 1882–1986, using the sequential clustering method.

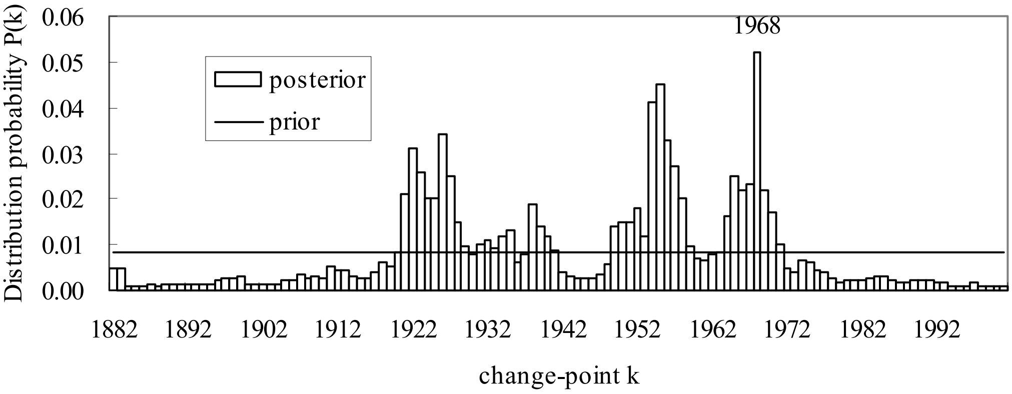

Fig. 7 The prior and posterior probability distribution of change-point location over the annual mean discharge series (1882–2001) of the Yangtze River.

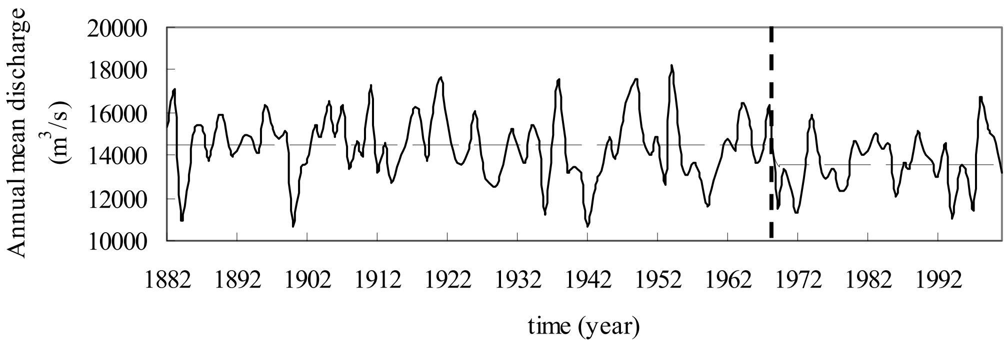

Fig. 8 The change of the mean levels over the annual mean discharge series (1882–2001) of the Yangtze River.

In order to further test whether or not the year 1993 is a potential change point for the annual mean series after 1968, the annual mean series of 1969–2001 was divided into two segments: 1969–1993, mean value of 13 578 $m^3 s^{-1}$, and 1994–2001, mean 13 654 $m^3 s^{-1}$. Again using the two-sample (two-tailed) *t*-test , the null hypothesis that the two segments separated by 1993 have the same mean level was not rejected at the below 5% significance level. That is, there is no evidence that the year 1993 is a change point in the annual mean series 1969–2001.

The change points for both the annual minimum and the annual mean series were found to have occurred before 1993, i.e. before the commencement of the Three Gorges Project. Hence, one can state that, so far, since the construction of the Three Gorges Project, there have not been any detectable significant changes in the stochastic characteristics of the time series of mean water level at the downstream Yichang station.

RELATIONSHIP BETWEEN GRADUAL CHANGE (TREND) AND ABRUPT CHANGE

Using the Bayesian change-point analysis, it was found that the mean levels of the annual minimum flow and the annual mean discharge series of the Yangtze River at the Yichang hydrological station are not constant, but decreased abruptly at some time. Both trend test methods and change-point detection methods have demonstrated that the annual minimum flow and the annual mean discharge have significantly decreased over the period 1882–2001, especially the former. As the trend term and the abrupt change (jump) term cannot be easily distinguished in statistical tests (Yevjevich, 1972), it is essential to check whether the time series have both a trend and an abrupt change, or just one of them. To investigate this interaction between the trend term and the abrupt change term, the following two-part analysis was carried out:

(a) The linear trend is abstracted from the original time series to generate a new time series. Then the change-point analysis methods are used to detect the most probable location of the change point in the new time series.

(b) The corresponding mean level is abstracted from each segment consisting of the original time series to generate a new time series. For example, to generate a new series from the annual minimum flow series, the mean value of 3707 $m^3 s^{-1}$ is deducted from each of the observed Q_{min} values from the period 1882–1934, while the mean value of 3408 $m^3 s^{-1}$ is deducted from each of the observed Q_{min} values from the period 1935–2001. Then the trend test methods are employed to check whether the new time series possess any linear trend.

By applying the above analyses to both the annual minimum flood and the annual mean discharge series, it was found that: (a) after the linear trend is deducted from each of the original series, the resultant two new series display no sign of a significant shift in the mean levels; and (b) after the corresponding mean values are deducted from each segment of the original time series, the resultant two new series display no sign of significant (linear) trend. In general, for the annual minimum and the annual mean discharge series of the Yangtze River at the Yichang hydrological station, both the trend and the abrupt change terms are very closely intertwined and it is very hard to distinguish between them.

TREND TEST FOR SOME NEW ANNUAL DISCHARGE SERIES

To make a more extensive assessment on Yangtze River discharge series, the Mann-Kendall, Spearman's rho and *t* tests were employed to test trends in some new annual discharge series extracted from the 120-year daily discharge series at the Yichang hydrological station. The new annual discharge series are established as follows. For each of the observed 365 (or 366) discharges in each year, a non-exceedence probability, denoted by p, is assigned to it according to its relative magnitude in that year. For example, if a discharge is larger than 100 discharges but less than 264 discharges in the same non-leap year, then its non-exceedence probability is calculated as $p = 100/365 = 27.4\%$. An annual series consisting of the daily discharge with the same non-exceedence probability p from each year is denoted by Q_p. The trend test results for the annual discharge series of different non-exceedence probability from $p = 0.05$ to 0.95 are listed in Table 2.

According to the results in Table 2, for the annual discharge series of $Q_{0.05}$, $Q_{0.10}$, $Q_{0.20}$, $Q_{0.30}$, $Q_{0.40}$, $Q_{0.50}$, $Q_{0.60}$ and $Q_{0.70}$, the null hypothesis of no trend can be rejected at the 5% significance level and so it is concluded that all these eight series have a decreasing trend with time. However, the null hypothesis of no trend cannot be rejected for the annual discharge series of $Q_{0.80}$, $Q_{0.90}$ and $Q_{0.95}$. These findings are

Table 2 Trend test of some new annual discharge series.

Series	$ac(1)$ *	Mann-Kendall test		Spearman's rho test		Linear regression	
		Z_{MK}	p value	Z_{SR}	p value	$\hat{b}$	T
$Q_{0.05}$	0.118	3.251	0.001	3.184	0.001	–3.600	–3.642
$Q_{0.10}$	0.122	2.357	0.018	2.271	0.023	–2.833	–2.693
$Q_{0.20}$	0.258	2.058	0.040	2.056	0.040	–2.996	–2.254
$Q_{0.30}$	0.201	2.169	0.030	2.172	0.030	–3.598	–2.067
$Q_{0.40}$	–0.053	2.409	0.016	2.398	0.016	–5.824	–2.170
$Q_{0.50}$	0.124	2.067	0.039	2.007	0.045	–9.308	–2.149
$Q_{0.60}$	0.240	3.260	0.001	3.339	0.001	–17.000	–3.233
$Q_{0.70}$	0.166	2.632	0.008	2.674	0.007	–17.713	–2.463
$Q_{0.80}$	0.065	2.185	0.029	2.182	0.029	–15.917	–1.744
$Q_{0.90}$	0.014	0.826	0.409	0.750	0.453	–6.987	–0.550
$Q_{0.95}$	0.054	0.798	0.425	0.752	0.452	–7.584	–0.549

* $ac(1)$: lag-1 autocorrelation coefficient of the series.

generally in agreement with the previous findings on the series of annual maximum, annual minimum and annual mean; that is, the null hypothesis of no trend can be rejected at the 5% significance level for both the annual minimum and annual mean series; however, it cannot be rejected at the 5% significance level for the annual maximum series. This agreement can be well expected as the series of annual minimum, annual mean and annual maximum are similar to the annual series of $Q_{0.05}$, $Q_{0.50}$ and $Q_{0.95}$, respectively, as far as the magnitude of the series is concerned.

DICUSSION AND CONCLUSIONS

The trend test and change-point analyses were carried out on three 120-year annual discharge series of the Yangtze River at the Yichang hydrological station: the annual maximum flood, the annual minimum flow and the annual mean discharge series for the period 1882–2001. The main purpose of the study was to check if these time series can really satisfy the consistency and stationarity assumptions for use in the future water resources planning and long-term reservoir operation of the Three Gorges Project.

Two widely used non-parametric methods, the Mann-Kendall test and the Spearman's rho test, were employed in the trend test of these three 120-year annual series. It was found that, at the 5% significance level, the annual maximum flood series did not show any statistically significant trend, while both the annual minimum flow series and the annual mean discharge series exhibited evidence of decreasing trend. The decreasing trend was most marked in the annual minimum flow series, and was attributed to the intensive water use for agriculture and industrial development in the upstream regions of the Yangtze River.

In the single change-point analysis, a Bayesian approach was adopted to analyse the shift of the mean levels in both the annual minimum flow series and the annual mean discharge series. For the annual minimum flow series, the most probable location of change point is the year 1934, after which the mean level decreased by nearly 300 $m^3 s^{-1}$ or 8%. For the annual mean discharge series, the most probable location of change point is the year 1968, after which the mean level decreased by nearly 900 $m^3 s^{-1}$ or 6%.

As the change points for both the annual minimum and the annual mean series occurred before 1993 (the year in which the Three Gorges Project commenced), one can state that, since the construction of the Three Gorges Project there have not been any significant changes in the annual minimum or the annual mean series.

However, it is very possible that the above conclusions might change with time, as the Three Gorges Project will definitely exert some influences on the flow regime of the Yangtze River at the Yichang hydrological station. Any change in the characteristics of the hydrological time series of Yichang station in the future could be a reason for modifying the initial construction and operation plan for the Three Gorges Project.

The consistency and stationarity of the stochastic characteristics of the 120-year hydrological time series at Yichang may be influenced by many factors, from different observation techniques to reservoir construction in the upstream region. However, as the catchment area above the Yichang station is very large (about 1 million km^2), more intensive and comprehensive study will be needed to foresee the influence of each factor on the future change of the hydrological time series at Yichang station.

Acknowledgements This research is supported by the Open Research Fund Program of State Key Laboratory of Water Resources and Hydropower Engineering Sciences (2003C002) and the Natural Sciences Fund of Hubei Province (2002AB009). The authors are very grateful to the editor and two referees for helpful comments on this paper.

REFERENCES

Adeloye, A. J. & Montaseri, M. (2002) Preliminary streamflow data analyses prior to water resources planning study. *Hydrol. Sci. J.* **47**(5), 679–691.

Berger, J. O. (1985) *Statistical Decision Theory and Bayesian Analysis*. Springer, New York, USA.

Box, G. E. P. & Jenkins, G. M. (1976) *Time Series Analysis: Forecasting and Control* (revised edn). Holden-Day, San Francisco, California, USA.

Buishand, T. A. (1984) Tests for detecting a shift in the mean of hydrological time series. *J. Hydrol.* **73**, 51–69.

Burn, D. H. (1994) Hydrologic effects of climatic changes in west central Canada. *J. Hydrol.* **160**, 53–70.

Burn, D. H. & Hag Elnur, M. A. (2002) Detection of hydrologic trends and variability. *J. Hydrol.* **225**, 107–122.

Changnon, S. A. & Kunkel, K. E. (1995) Climate-related fluctuation in Midwestern floods during 1921–1985. *J. Water Resour. Plan. Manage.* **121**(4), 326–334.

Chernoff, H. & Zacks, E. (1963) Estimating the current mean of a normal distribution which is subjected to changes in time. *Ann. Math. Statist.* **35**, 1999–1028.

Chib, S. (1996) Calculating posterior distributions and modal estimates in Markov mixture models. *J. Econometrics* **75**, 79–98.

Chib, S. (1998) Estimation and comparison of multiple change-point models. *J. Econometrics* **86**, 221–241.

Crawshaw, J. & Chambers, J. (1990) *A Concise Course in A-Level Statistics with Worked Examples.* Stanley Thornes Ltd., Avon, UK.

Douglas, E. M., Vogel, R. M. & Kroll, C. N. (2000) Trends in floods and low flows in the United States: impact of spatial correlation. *J. Hydrol.* **240**, 90–105.

Gan, T. Y. (1998) Hydroclimatic trends and possible climatic warming in the Canadian Prairies. *Water Resour. Res.* **34**(11), 3009–3015.

Hirsch, R. M. & Slack, J. R. (1984) Non-parametric trend test for seasonal data with series dependence. *Water Resour. Res.* **20**(6), 727–732.

Hubert, P., Carbonnel, J. P. & Chaouche, A. (1989) Segmentation des séries hydrométriques. Application à des séries de précipitations et de débits de l'Afrique de l'Ouest. *J. Hydrol.* **110**, 349–367.

Kendall, M. G. (1975) *Rank Correlation Methods*. Griffin, London, UK.

Kotz, S. & Wu, X. Z. (2000) *Modern Bayesian Statistics* (in Chinese). China Statistics Press, Beijing, China.

Kuczera, G. & Parent, E. (1998) Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. *J. Hydrol.* **211**, 69–85.

Lavielle, M. & Lebarbier, E. (2001) An application of MCMC methods for the multiple change-points problem. *Signal Processing* **81**, 39–53.

Lee, A. S. F. & Heghinian, S. M. (1977) A shift of the mean level in a sequence of independent normal random variables—a Bayesian approach. *Technometrics* **19**, 503–506.

Lehmann, E. L. (1975) *Nonparametrics, Statistical Methods Based on Ranks*. Holden-Day Inc., San Francisco, California, USA.

Lettenmaier, D. P., Wood, E. F. & Wallis, J. R. (1994) Hydro-climatological trends in the continental United States, 1948–88. *J. Climate* **7**, 586–607.

Lins, H. F. & Slack, J. R. (1999) Streamflow trends in the United States. *Geophys. Res. Lett.* **26**, 227–230.

Mann, H. B. (1945) Nonparametric tests against trend. *Econometrica* **13**, 245–259.

Meddis, R. (1984) *Statistics using Ranks, A Unified Approach*. Blackwell, Oxford, UK.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. (1953) Equations of state calculations by fast computing machines. *J. Chem. Phys.* **21**, 1087–1092.

Olsen, J. R., Stedinger, J. R., Matalas, N. C. & Stakhiv, E. Z. (1999) Climate variability and flood frequency estimation for the upper Mississippi and lower Missouri rivers. *J. Am. Water Resour. Assoc.* **35**(6), 1509–1520.

Paturel, J. E., Servat, E., Kouamé, B., Lubès, H., Quedraogo, M. & Mason, J. M. (1997) Climatic variability in humid Africa along the Gulf of Guinea. Part II: An integrated regional approach. *J. Hydrol.* **191**, 16–36.

Perreault, L., Haché, M., Slivitsky, M. & Bobée, B. (1999) Detection of changes in precipitation and runoff over eastern Canada and US using a Bayesian approach. *Stochastic Environ. Res. Risk Assessment* **13**, 201–216.

Perreault, L., Bernier, J., Bobée, B. & Parent, E. (2000a) Bayesian change-point analysis in hydrometerological time series. Part 1: The normal model revisited. *J. Hydrol.* **235**, 221–241.

Perreault, L., Bernier, J., Bobée, B. & Parent, E. (2000b) Bayesian change-point analysis in hydrometerological time series. Part 2: Comparison of change-point models and forecasting. *J. Hydrol.* **235**, 242–263.

Pupacko, A. (1993) Variations in northern Sierra Nevada streamflow: implications of climate change. *Water Resour. Bull.* **29**(2), 283–290.

Qin, A. J., Chen, X. Y. & Zhen, Y. X. (1993) The stochastic analysis of the discharge series at Yichang. *Chinese J. Hydrol.* **13**(5), 15–22.

Servat, E., Paturel, J. E., Lubès, H., Kouamé, B., Quedraogo, M. & Mason, J. M. (1997) Climatic variability in humid Africa along the Gulf of Guinea. Part I: detailed analysis of the phenomenon in Côte d'Ivoire. *J. Hydrol.* **191**, 1–15.

Smith, A. F. M. & Roberts, G. O. (1993) Bayesian computation via Gibbs sampler and related Markov Chain Monte Carlo methods. *J. Roy. Statist. Soc.* (B) **55**(1), 3–23.

Smith, K. & Richman, M. B. (1993) Recent hydroclimatic fluctuations and their effects on water resources in Illinois. *Climatic Change* **24**, 249–269.

Sneyers, R. (1990) On the statistical analysis of series of observations. Tech. note no. 143, WMO no. 415, World Meteorological Organization, Geneva, Switzerland.

Utts, J. M. & Heckard, R. F. (2002) *Mind on Statistics*. China Machine Press, Beijing, China.

Xiong, L. H. & O'Connor, K. M. (2000) Analysis of the response surface of the objective function by the optimum parameter curve: How good can the optimum parameter values be? *J. Hydrol.* **234**, 187–207.

Yevjevich, V. H. (1972) *Stochastic Processes in Hydrology.* Water Resources Publications, Littleton, Colorado, USA.

Yu, J. R., Tzeng, G. H. & Li, H. L. (2001) General fuzzy piecewise regression analysis with automatic change-point detection. *Fuzzy Sets and Systems* **119**, 247–257.

Yue, S. & Wang, C. Y. (2002) Applicability of pre-whitening to eliminate the influence of serial correlation on the Mann-Kendall test. *Water Resour. Res.* **38**(6), 10.1029/2001WR000861.

Yue, S., Pilon, P. & Cavadias, G. (2002a) Power of the Mann-Kendall and Spearman's rho test for detecting monotonic trends in hydrological series. *J. Hydrol.* **259**, 254–271.

Yue, S., Pilon, P., Phinney, B. & Cavadias, G. (2002b) The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrol. Processes* **16**(9), 1807–1829.

Zhan, D. J. & Ye, S. Z. (2000) *Engineering Hydrology* (in Chinese). Chinese Water Power Press, Beijing, China.

Zhang, X., Harvey, K. D., Hogg, W. D. & Yuzyk, T. R. (2001) Trends in Canadian streamflow. *Water Resour. Res.* **37**(4), 987–998.

Received 16 April 2003; accepted 22 October 2003